DOCUMENT RESUME

ED 097 337 TM 003 898

| | |
|---|---|
| AUTHOR | Sadacca, Robert; Brackett, Joan |
| TITLE | The Validity and Discriminatory Impact of the Federal Service Entrance Examination. |
| INSTITUTION | Urban Inst., Washington, D.C. |
| SPONS AGENCY | Ford Foundation, New York, N.Y. |
| PUB DATE | Sep 71 |
| NOTE | 32p. |
| AVAILABLE FROM | Publications Office, The Urban Institute, 2100 M St. N.W., Washington, D. C. 20037 ($1.50) |
| EDRS PRICE | MF-$0.75 HC-$1.85 PLUS POSTAGE |
| DESCRIPTORS | Comparative Analysis; *Government Employees; Negroes; Personnel Selection; *Racial Discrimination; *Screening Tests; *Test Bias; *Test Validity |
| IDENTIFIERS | *Federal Service Entrance Examination |

ABSTRACT

The Federal Service Entrance Examination (FSEE) is one of the competitive examinations used to screen junior level applicants for managerial, technical, and professional occupations in some 50 U. S. Government agencies. It consists of 95 objective test questions designed to measure the applicant's vocabulary, reading comprehension, knowledge of English usage, and ability to reason quantitatively in solving problems with numbers. The objective of this research is to provide information relevant to two questions: (1) Is there sufficient evidence that the FSEE is screening out applicants who would have done more poorly in Federal service than those who pass the test to warrant the continued use of the FSEE? and (2) Do proportionately fewer black applicants pass the FSEE than white applicants? The validity of the FSEE is explored, and the results of analyses of data obtained from the Civil Service Commission are presented. FSEE passing rates are compared for students from black and white colleges. It is concluded that the use of the FSEE is unfairly discriminatory to many black applicants, and its use should be suspended pending strong evidence of its validity. (Author/RC)

ED 097337

# THE VALIDITY AND DISCRIMINATORY IMPACT OF THE FEDERAL SERVICE ENTRANCE EXAMINATION

ROBERT SADACCA
assisted by
Joan Brackett

712-2-1 September 1971

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM 003 898

THE URBAN INSTITUTE
WASHINGTON, D.C.

The research and studies forming the basis for this publication were conducted with financial support of the Ford Foundation. The views expressed are those of the authors and do not necessarily represent the views of the sponsor of the research or of The Urban Institute.

UI 712-2-1

Available from

Publications Office
The Urban Institute
2100 M Street N.W.
Washington, D.C. 20037

List price: $1.50

# CONTENTS


| | Page |
|---|---|
| PREFACE | ii |
| I. INTRODUCTION | 1 |
| II. THE VALIDITY OF THE FSEE | 5 |
| A. Types of Validity | 5 |
| B. Standards and Guidelines for Validation | 6 |
| C. Available Validity Data | 9 |
| D. Appraisal of the Validity of the FSEE | 11 |
| III. THE DISCRIMINATORY IMPACT OF THE FSEE | 14 |
| A. The Meaning of Test Unfairness | 14 |
| B. Comparison of Black and White FSEE Passing Rates | 15 |
| IV. CONCLUSION | 22 |
| APPENDIX A | 23 |
| APPENDIX B | 26 |
| APPENDIX C | 27 |

# PREFACE

The Urban Institute is undertaking studies in a number of fields where the expertise and skills of the staff coincide with issues in current litigation or administrative proceedings. This paper on the Federal Service Entrance Examination represents one such effort. Topics for studies are suggested by public officials, lawyers, or interested Institute staff members. Regardless of the source of a research issue, all resulting publications represent the objective and policy-oriented approach of The Urban Institute. The findings and conclusions of the studies will be available to judges, administrators, lawyers, and to the public generally.

This report was prepared by Robert Sadacca with the assistance of Joan Brackett. Mr. Sadacca holds a Ph.D. in Psychometrics from Princeton University and is a senior staff member at The Urban Institute. He was a Psychometric Fellow at the Educational Testing Service, Executive Editor of the Psychological Abstracts, and a Program Director at the U.S. Army Behavioral Science Research Laboratory. Mrs. Brackett is on the federal evaluation practices staff at the Institute.

# I. INTRODUCTION

The Federal Service Entrance Examination (FSEE) is produced by the Civil Service Commission under governing legislation which calls for applicants for federal employment to be evaluated through open, competitive examinations which "are practical in character and as far as possible relate to matters that fairly test the relative capacity and fitness of the applicants for the appointments sought."[1] The examination is further produced under the mandate "to insure equal employment opportunities for employees without discrimination because of race, color, religion, sex or national origin."[2] This study of the validity and possible discriminatory impact of the FSEE has been undertaken by The Urban Institute in order to help those concerned with the development, application and use of this selection instrument to evaluate how well the Civil Service Commission is discharging its responsibility.

The Federal Service Entrance Examination (FSEE) is one of some 20 competitive examinations produced by the Commission. It is currently used to screen junior level applicants for approximately 200 managerial, technical and professional occupations in some 50 U.S. Government agencies. The examination consists of 95 objective test questions designed to measure the applicant's vocabulary, reading comprehension, knowledge of English usage, and ability to reason quantitatively in solving problems with numbers.

---

[1] 5 U.S.C. §3301; 5 U.S.C §3304

[2] 5 U.S.C. §7151

The difficulty level of the questions is pitched for college seniors and graduates. Cut-off scores based on a minimum number of correct answers are set on the basis of normative data and projected man-power needs and are applied across the 200 occupations--those applicants scoring at or above the cut-off points are considered eligible for Federal Civil Service employment; those scoring below the cut-off points are generally considered ineligible. Different cut-off scores are set for entry into the Federal Civil Service at the GS-5, GS-7 and GS-9 levels.

Passing the FSEE does not insure the appointment of an applicant. In 1969, of the approximately 101,000 individuals who took the FSEE, 49,500 passed, but only about 7,700 received appointments. The names of applicants who pass the FSEE are placed in rank order on the basis of their scores on Federal Registers for the particular occupation or program for which the individual applied. As appropriate vacancies develop in the various federal agencies, the FSEE scores, education and job experience of the applicants are appraised against the requirements of the agency positions. The emphasis placed on the FSEE in the final appraisal of an applicant varies from agency to agency and within each agency varies across the job spectrum. However, in general, the higher his FSEE score, the greater the probability an applicant will receive an appointment.

The FSEE has been used substantially in the same manner since 1955 for screening college-trained applicants. During the 16 years it has been used as a screening instrument, the type of questions that make up the test have changed somewhat: items measuring spatial visualization ability were deleted about ten years ago and relatively greater emphasis was placed on assessing verbal rather than quantitative ability. Though

applicants to some occupations no longer have to take the FSEE, the number of occupations for which a passing FSEE score is required has increased since its initial use.

Persons are sometimes hired even though they have not taken the FSEE. To ascertain the extent of its use at regional and state levels, the Institute surveyed HEW, DOL, OEO, and HUD Regional Offices and CSC Regional and Area Offices. Survey results (see Appendix A) showed the pattern of use of the FSEE to be a mixed one for HEW, OEO and HUD. Since 1967, in order to facilitate greater minority group placement in managerial and administrative positions, the FSEE has not been required if the applicant has earned a bachelor's degree within the previous two years, and maintained at least a 3.5 (B+) average or ranked in the top ten percent of his class. (Approximately five percent of those appointed from the FSEE Register enter the Federal Service through these means each year).

The use and non-use of the FSEE raise many questions concerning the appropriateness and fairness of applying the FSEE as a screening and selective instrument in federal employment. The FSEE essentially measures only two aptitudes: verbal and quantitative abilities. However practical the FSEE may be to administer, considering the myriad positions in government agencies and the uniqueness and many talents of the thousands of individuals seeking these positions, it is questionable whether the use of cut-off scores based on one 95-item test to determine eligibility for federal employment is consistent with the Federal statutes cited earlier.

A comprehensive examination of all of the issues involved in the use of mass screening instruments like the FSEE is beyond the scope

of this Urban Institute effort. The limited objective of this research is to provide information relevant to two questions:

(1) Is there sufficient evidence that the FSEE is screening out applicants who would have done more poorly in the Federal Service than those who pass the test to warrant the continued use of the FSEE?

(2) Do proportionately fewer black applicants pass the FSEE than white applicants?

The next section explores the validity of the FSEE; that is, the relationship between scores on that test and indexes of selected applicants' performance on the jobs for which the FSEE is used as a screening instrument. Section III of this report presents the results of analyses of data obtained from the Civil Service Commission. FSEE passing rates are compared for students from black and white colleges. The last section presents our conclusion concerning whether the FSEE is unfairly discriminatory.

## II. THE VALIDITY OF THE FSEE

### A. Types of Validity

The validity of a test refers to the relationships between scores achieved by individuals taking the test and the job performance of the individuals. Three types of validity are generally recognized:[3]

(1) Content Validity--the extent to which the test content is a representative sample of the universe of tasks, processes and conditions constituting the job. The aptitudes, skills and knowledge required to achieve good test performance are judged by experts to be the aptitudes, skills and knowledge required for successful job performance. A typing test administered to secretarial applicants is an example of a test having some content validity.

(2) Criterion Validity--the degree of relationship or correlation between test scores and a criterion, that is, a variable considered to be a direct measure or index of job performance. The correlation between scores on the College Board Entrance Examination and graduation/non-graduation from a college is an example of criterion validity.

(3) Construct Validity--the interpretation of a test as a measure of a theoretical variable or construct that is presumed to be important for successful performance in a job or set of jobs. The construct may be a general ability, trait or attitude that can be said to underlie

---

[3] See: Standards for Educational and Psychological Tests and Manual, American Psychological Association, Washington, D.C. 1966.

Federal Register, Volume 35, Title 29--Labor, Chapter XIV--Equal Employment Opportunity Commission, Part 1607--Guidelines on Employment Selection Procedures, August 1, 1970, p. 3.

performance in a number of tasks or jobs. A test measuring achievement motivation for salesmen is an example of construct validity.

B. Standards and Guidelines for Validation

Tests are frequently used in arriving at decisions which may have great influence on the ultimate welfare of the persons tested, on educational and employment points of view and practices, and on the development and utilization of human resources. The more valid the test, the greater the likelihood that high caliber personnel will be selected to fill available positions. Test users, therefore, need to apply high standards of professional judgment in selecting and interpreting tests, and test producers are under obligation to produce tests which can be shown to be valid.

Standards and guidelines for validation have been established and are accepted and utilized in the employment testing field, particularly in private employment.[4] Involved in any concern with validity should be information about initial recruiting practices; the conditions and procedures of test administration; distribution of scores by major types of applicants and selected candidates; how the test scores are being used in the selection process; the reliability of the test (a measure of the stability of test scores over time); the relationship of test scores to scores on other tests; job analyses and descriptions; and performance evaluation and promotion procedures. This information should be utilized in a methodologically acceptable validation study. It should be emphasized

---

[4] Federal Register, Ibid. p. 5.

that as jobs and applicants change over time, test items can become outdated or inappropriate. The validity of a test must be periodically reassessed.

Content validity is particularly subject to change as it is most dependent upon job tasks remaining constant. Where content validity is claimed, it is important to carefully examine the comparative analyses of the test and job in order to assess the degree of overlap or representativeness of the test tasks and conditions to the job tasks and conditions. As considerable judgment is involved, the professional experience and qualifications of the experts who made the validity statements should be taken into consideration. Evidence of the extent of agreement among the experts should also be considered.

In assessing criterion validity particular attention must be paid to the criterion itself. Does the index truly measure job performance? Are there significant aspects of performance that the criterion measure does not embrace or reflect? How long after administration of the test are the criterion measures obtained and how reliable are they? Was knowledge of test scores allowed to affect the performance index assigned to the individuals? Is performance assessment unbiased by such factors as the education, age, sex and race of the individuals or of the judges making the assessments?

These questions should also be kept in mind in construct validity even though actual criterion measures are not obtained. Anastasi[5] states:

---

[5]Anna Anastasi, "Some Current Developments on the Measurement and Interpretation of Test Validity," in Testing Problems in Perspective, Anna Anastasi, ed., American Council on Education, Washington, D.C. 1966, p. 309.

> Actually the theoretical construct or trait assessed by any test can be defined in terms of the operations performed in establishing the validity of the test. Such a definition should take into account the various external criteria with which the test is correlated significantly, as well as the conditions that affect its scores. It would also seem desirable to retain the concept of criterion in construct validation, not as a specific practical achievement to be predicted, but as a general name for independently gathered external data. The need to base all validation on data rather than on armchair speculation would thus be re-emphasized, as would the need for data external to the scores themselves.

The use of construct validity in prediction of job performance requires that a test correlate highly with other variables with which it should correlate and not correlate with variables with which it should differ. That is, the interpretation of the underlying theory must be substantiated and the evidence reported that constructs other than those assumed do not account for variances in scores on the test. It calls for constant accumulation of data on the nature of the trait under consideration and the conditions affecting its development and manifestations.

The use of a variety of data gathering techniques causes need for caution in assessing construct validity. One must be aware of "... highly opportunistic selection of evidence and the editorial device of failing to mention validity probes that were not confirmatory."[6] "Some test constructors apparently interpret construct validation to mean content validity expressed in terms of psychological trait names. Hence

---

[6] J. T. Campbell, "Recommendations for APA Test Standards Regarding Construct, Trait, or Discriminant Validity," American Psychologist, 1960, 15, p. 551.

they present as construct validity purely subjective accounts of what they believe (or hope) the test measures."[7]

In summary, test producers have the task of providing information about each test so that all concerned will know what reliance can be placed on it; presentation of validity evidence must be available to test users, and to the public.[8] Under no circumstances should the general reputation of a test, its authors, or casual reports of test utility or other non-empirical or anecdotal accounts of testing practices or testing outcomes be accepted in lieu of evidence of validity.[9]

C. Available Validity Data

All attempts to find any comprehensive published studies showing the validity of FSEE scores have been futile. Inquiries[10] at the Civil Service Commission produced the information that occasional studies for selected positions have been initiated from time to time by the Commission or some government agency. These studies, however, are not available to the public for the following reasons:

---

[7] Anna Anastasi, "Some Current Developments on the Measurement and Interpretation of Test Validity," in Testing Problems in Perspective, Anna Anastasi, ed., American Council on Education, Washington, D.C., 1966, p. 309.

[8] Standards for Educational and Psychological Tests and Manuals, American Psychological Association, Washington, D.C., 1966, p. 1.

[9] Federal Register, Volume 35, Title 29, Labor, Chapter XIV--Equal Employment Opportunity Commission, Part 1607--Guidelines on Employment Selection Procedures, August 1, 1970, p. 3.

[10] Interviews with Mr. Herbert Ozur, Bureau of Research and Examinations, December 10, 1970, and Dr. Albert Maslow, Chief, Personnel Measurement and Research, Civil Service Commission, December 18, 1970.

1. They are not sufficiently rigorous scientifically to be allowed public appraisal.

2. At the time of the Civil Service Commission move in 1964, many of the records were destroyed; the remainder are at the Records Center and it would be an impossible task to locate them.

3. The research done on the FSEE for the agencies is located in memos in scattered offices of the persons concerned at the time. Furthermore, it is not known if the individual agencies would release whatever research they may have on FSEE.

Research at the Civil Service Commission Library located descriptive summary statistics of the competitor populations,[11] some general discussions of the rationale of the use of the FSEE,[12] and some material on performance ratings and job analysis. The one or two possible relevant articles[13] listed in the Bibliography of Civil Service Publications were not catalogued or could not be found by the library staff. Two possibly relevant reports[14] were catalogued CONFIDENTIAL.

---

[11] Description of the Competitive Population: Federal Service Entrance Examination, August, 1957 and January 1958, Assembled Test Technical Series No. 30, U.S. Civil Service Commission, Washington, D.C.

[12] Rationale for the FSEE and Considerations in the Use of the FSEE Written Test in Occupation Selection. (Ie 179 Un35r)

[13] Validity Study of Engineering Aid Examination, Unassembled Rating Procedure No. 1 and Validity Study for Semi-Automatic Teletype Repairmen, November, 1947. (Bibliography of Civil Service Publications)

[14] A Statistical Report on the FSEE, 1956 and Federal Service Entrance Examination Management Intern Interviewing and Rating Procedures.

None of these materials, however, could be considered serious efforts at establishing the validity of a test being used to screen 100,000 applicants annually for 200 job classifications. It is our understanding that Commission staff members are currently conducting an in-house study of seven job positions using work satisfaction, specially developed performance ratings, FSEE and other test scores, and a number of other variables. Another study[15] is being conducted to determine the interrelationships of FSEE and other test scores with performance in Defense Department inventory management positions. The results of these studies have not been published to date, and could, at best, validate a small proportion of the use to which the FSEE is put.

The absence of a comprehensive validity study of the FSEE is not surprising considering the formidable scope of such an effort. The methodological problems involved in obtaining the necessary data would be enormous because of the numerous positions involved, the non-uniform manner in which applicants are hired by agencies, the high turnover rate of employees, and the dependence of promotions and ratings on the unique circumstances within the individual agencies. It is not our intent to suggest that such a study should be undertaken at this time. We wish simply to indicate that there is no available published evidence that the FSEE has been validated in accordance with generally accepted standards and guidelines for employment tests.

D. Appraisal of the Validity of the FSEE

In the absence of the necessary supportive data any appraisal of the validity of the FSEE must be purely conjectural. Officials at the Civil

---

[15] Being conducted for the Civil Service Commission by the Educational Testing Service, Princeton, New Jersey.

THIS PAGE WAS MISSING FROM THE DOCUMENT THAT WAS SUBMITTED TO ERIC DOCUMENT REPRODUCTION SERVICE.

pages 12 and 13

## III. THE DISCRIMINATORY IMPACT OF THE FSEE

### A. The Meaning of Test Unfairness

Flaugher[18] distinguishes three sources of unfairness existing within some testing practices:

> The first and by far the most commonly referred to is that of the test content. There is a widely held belief that the kinds of questions asked within the test, are biased against minority groups, causing them to perform poorly in ways that are not valid. Second, the test program itself may be conducted in such a way that the result is discriminatory. For example, information essential to registering for and taking the test may not be disseminated in a form that makes it available to minority groups, or conditions may be allowed to exist in the test administration itself which are intimidating. Third, discriminatory practices may exist in the use to which test results are put, such as requiring high verbal test scores to qualify for a job which in fact does not depend upon verbal skills, . . .

Flaugher emphasizes that "Lower test scores for a particular minority group are not in themselves evidence of unfair testing practices. Assuming a fair, unbiased criterion, ultimate conclusions about the question of test bias must rest on evidence concerning the validity of the particular tests in predicting the criterion."[19]

It is our contention that since the validity of the FSEE has not been adequately demonstrated, if it can be shown that a minority group,

---

[18] Ronald L. Flaugher, "Testing Practices, Minority Groups, and Higher Education: A Review and Discussion of the Research," Research Bulletin 70-41, Educational Testing Service, Princeton, New Jersey, June 1970, pp. 6-7.

[19] Ibid., p. 11.

specifically blacks, performs in general more poorly on the FSEE and is subsequently disproportionately barred from federal employment, then the use of the FSEE is unfairly discriminatory. As the reasons underlying black/white differences in test performance are immaterial to this argument, they will not be discussed here.[20]

B. Comparison of Black and White FSEE Passing Rates

As the Civil Service Commission does not require applicants to submit information concerning their race, a comprehensive statistical comparison of the performance of black and white applicants on the FSEE is not possible from existing data. However, The Urban Institute was able to obtain from the Civil Service Commission statistics showing the number of college seniors and graduate students who took the FSEE and the number who passed the FSEE at a large number of universities and colleges in June 1968 and during the 1968-1969 school year. From the Office of Civil Rights, HEW, The Urban Institute obtained the number of black undergraduates and the number of black graduate students, the total number of undergraduate and the total number of graduate students at many universities and colleges for

---

[20] A vast literature exists concerning the difference in test results between minority groups and whites and the reasons underlying these differences. The interested reader is referred to:

A. Shuey: The Testing of Negro Intelligence (2nd ed.), New York, Social Science Press, 1966.

Clay L. Moore, Jr., John F. MacNaughton, and Hobart G. Osburn, "Ethnic Differences Within an Industrial Selection Battery," Personnel Psychology, Volume 22, 1969.

Howard C. Lockwood, "Critical Problems in Achieving Equal Employment Opportunity," Personnel Psychology, Spring 1966.

Gordon Fifer, "Social Class and Cultural Group Differences in Diverse Mental Abilities," Testing Problems in Perspective, ed. by Anna Anastasi, American Council on Education, Washington, D.C., 1966.

Fall 1968. By comparing the lists of universities and colleges obtained from the Civil Service Commission and the Office for Civil Rights, a subset of institutions was obtained for which both racial enrollment and applicant FSEE performance were available.

From this data there was, of course, still no way to tell the race of the applicants. However, by eliminating all institutions where the percentage black was greater than 1 percent but less than 99 percent, two subgroups of institutions became available: one consisting of institutions with 99 percent or more black enrollment and the other consisting of institutions with 99 percent or more white enrollment. Recognizing that there would be some classification error, it seemed reasonable to assume for the purpose of further analysis that all applicants in these schools were either black or white.

In order to control for differences in the educational background of the applicants, each of 50 black schools was paired with a matching white school. Using information from the data bank of the American Council on Education,[21] the pairing of the black with the white schools was made on the following factors: location (pairs had to be within the same state or adjoining state), number of students, number of professors with Ph.D.'s, endowment per student, and type of school (private or public; co-ed, all male, all female). These particular factors were selected owing to their availability and our judgment that they might be related to the educational opportunities available to the student. Table 1 shows the sample of paired schools by state while Table 2 shows the distribution of the schools on three of

---

[21] John A. Creager and Charles L. Sell, The Institutional Domain of Higher Education: A Characteristics File for Research, ACE Research Reports, Volume 4, American Council on Education, 1969, p. 6.

TABLE 1

SAMPLE OF MATCHED BLACK AND WHITE SCHOOLS BROKEN DOWN BY STATE

| State | Black | White | Total |
|---|---|---|---|
| Alabama | 5 | 6 | 11 |
| Arkansas | 1 | 1 | 2 |
| Florida | 4 | 3 | 7 |
| Georgia | 6 | 7 | 13 |
| Louisana | 4 | 4 | 8 |
| Maryland | 1 | 1 | 2 |
| Mississippi | 4 | 3 | 7 |
| North Carolina | 7 | 7 | 14 |
| Oklahoma | 1 | 1 | 2 |
| South Carolina | 4 | 3 | 7 |
| Tennessee | 5 | 6 | 11 |
| Texas | 7 | 7 | 14 |
| Virginia | 1 | 1 | 2 |
| TOTAL | 50 | 50 | 100 |

the matching factors. All but three of the schools were co-ed; 31 of the black schools were private compared to 35 of the white schools.

TABLE 2

MEAN NUMBER OF STUDENTS, PROFESSORS WITH Ph.D.'s AND ENDOWMENT PER STUDENT IN SAMPLE OF MATCHED BLACK AND WHITE SCHOOLS

| | Black | White |
|---|---|---|
| Number of Students | 1,489 | 1,475 |
| Number of Ph.D. Professors | 19.4 | 19.6 |
| Endowment per Student ($) | 1922 | 1875 |

Estimates of the percentage of seniors at the black and the white schools who elected to take the FSEE were obtained in order to determine whether there were wide differences in application rates between the sets of schools. (Significant differences in application rates would indicate that the Federal Service might be appealing to different segments of the white and black student bodies.) The estimates obtained are quite rough because the number of seniors at the schools was estimated through dividing the number of undergraduates by four. Estimates of the percentages of graduate students at the schools who elected to take the FSEE were also obtained. Table 3 contains the mean of these estimated percentages for the black and white schools. The percentages are presented here not because they are accurate estimates of the average application rates at the schools, but to show that the application rates at both the black and white schools were low and fairly comparable in value. Altogether 984 senior and graduate

students at the black schools took the FSEE compared to 924 students at the white schools.

TABLE 3

MEAN ESTIMATED APPLICATION RATES (PERCENTAGES) IN SAMPLE OF MATCHED BLACK AND WHITE SCHOOLS

| | Black | White | No. of Paired Schools in Average* |
|---|---|---|---|
| Seniors | | | |
| June 1968 | 3.4 | 2.2 | 40 |
| 1968-1969 | 1.9 | 1.6 | 44 |
| Graduate Students | | | |
| June 1968 | 8.9 | 2.3 | 4 |
| 1968-1969 | 7.0 | 8.8 | 4 |

* The number of paired schools does not equal 50 owing to the absence of FSEE data for the time period indicated.

The rate at which the applicants passed the FSEE was compared in two ways. First, the passing rates at each pair of schools were compared to find out how many times the rate at the black schools was greater than the rate at their paired white schools and vice versa. Sign tests[22] were then

[22] Sign tests were used in order to avoid any assumptions concerning the form of the distribution of passing rate differences. Chi-square, similarly, is a non-parametric or distribution-free test. See Chapters 8 and 9 of Dean J. Champion, Basic Statistics for Social Research, Chandler Publishing Company, Scranton, Pennsylvania, 1970, for a discussion of sign-tests and Chi-square.

employed to statistically evaluate the significance of the median of the differences in rates obtained. Second, the available data from the matched pairs was pooled across the black and white schools, that is, the data from all black schools and all white schools were combined separately. Chi-square tests were then run to test the significance of the overall pass rates obtained for the black and white students taking the FSEE. Table 4 shows the median pass rates at the black and white schools and the results of the sign tests. Table 5 shows the pooled pass rates and the results of the Chi-square tests.

The data in Tables 4 and 5 clearly show that the FSEE pass rates of black students is generally lower than that of white students at comparable schools. The sign tests indicate that the hypothesis that the median pass rates among seniors are the same in the black and white schools must be rejected. Similarly the Chi-square tests indicate that the passing rates of each group of black students taking the FSEE was significantly lower than the rates of the white students. Altogether 8.6 percent of the black students passed the FSEE compared to 42.1 percent of the white students. We do not feel it is necessary to explain why these differences occur, although they are certainly consonant with racial differences in aptitude test performance obtained in many other studies.[23] It is sufficient to point out that a smaller percentage of blacks than whites are considered "qualified" to enter Federal Service at the GS-5 to GS-9 level for 200 job classifications as the result of the administration of the FSEE and that the validity of the FSEE as a screening instrument has not been satisfactorily demonstrated.

---

[23] Op. cit. Shuey and others

TABLE 4

MEDIAN FSEE PASS RATES (PERCENTAGES) IN SAMPLE OF MATCHED BLACK AND WHITE SCHOOLS AND RESULTS OF SIGN TESTS

| Year | Median Pass Rates | | No. of Schools Rate Greater | | | |
|---|---|---|---|---|---|---|
| | Black | White | B over W | W over B | Tied | Significance |
| Seniors | | | | | | |
| 1968 | 19.0 | 52.0 | 1 | 39 | 0 | P<.001 |
| 1969 | 26.7 | 53.7 | 5 | 33 | 6 | P<.001 |
| Graduate Students | | | | | | |
| 1968 | 11.1 | 41.7 | 0 | 3 | 1 | * |
| 1969 | 8.1 | 83.3 | 1 | 2 | 1 | * |

* There are an insufficient number of matched schools with graduate student data for the sign test to be realistically employed.

TABLE 5

PASS RATES (PERCENTAGES) OF BLACK AND WHITE FSEE EXAMINEES FROM SAMPLE OF MATCHED SCHOOLS AND RESULTS OF CHI-SQUARE TESTS

| Year | Black Examinees | | | White Examinees | | | |
|---|---|---|---|---|---|---|---|
| | No. Taking | No. Passing | Pass Rate | No. Taking | No. Passing | Pass Rate | Significance |
| Seniors | | | | | | | |
| 1968 | 537 | 36 | 6.7 | 433 | 202 | 46.6 | P<.001 |
| 1969 | 316 | 40 | 12.7 | 446 | 172 | 38.6 | P<.001 |
| Graduate Students | | | | | | | |
| 1968 | 65 | 4 | 6.2 | 21 | 7 | 33.3 | P<.01 |
| 1969 | 66 | 5 | 7.6 | 24 | 8 | 33.3 | P<.01 |

## IV. CONCLUSION

After examination of the available information concerning the validity of the FSEE and the comparative passing rates of black and white students from matched schools, the conclusion that the use of the FSEE is unfairly discriminatory to many black applicants seems unavoidable. It is unfortunate that this test has been used extensively for so many years with apparent inadequate validation. Pending strong evidence of its validity, the operational use of the FSEE for screening applicants to Federal Service should be suspended.

# APPENDIX A

In order to ascertain the extent and kind of use to which the FSEE is put in regional employing offices, The Urban Institute sent questionnaires to the following offices:

HUD, OEO, HEW, and DOL Regional Offices
CSC Regional Offices
CSC Area Offices (Federal Job Information Centers)

A copy of the letter and the survey instrument follow. (Appendices B and C) The number of questionnaires returned may be seen in Table A-1.

Of those replying to the question, "To what extent are personnel initially drawn from the FSEE register?", 58 percent reported that recruitment is mostly through their agency; 25 percent recruit mostly from the FSEE register; 16 percent recruit through their agency exclusively with no contact with the FSEE. One-third of the respondents reported giving the FSEE no weight in the total evaluation of the applicant, 22 percent give it only 10 percent, one-third give it 50 percent, and only 13 percent give it as much as 75 percent. The weight given varies with the applicant in 25 percent of the responses and with the position in 13 percent; the remainder report no variance in the weight given the FSEE in the total evaluation of the applicant. One interesting HUD response: "Under The Urban Intern Program, we are mostly concerned with the potentialities and capabilities of the individual rather than with the ranking made in the FSEE. Candidates' capacity for learning and development is tantamount in the selection of the applicant."

The data available varied considerably. Forty percent had either biographical data or FSEE scores broken down by testing site, 40 percent

had both, and 10 percent had just the FSEE scores. (Ten percent did not know.) Curiously, no offices had data by minority and/or ethnic group on "those made eligible by the FSEE and presently working." No office requested or participated in any study of the validity of the FSEE. However, one reply from an HEW regional office offers comments: "We have hired at the entrance level students in the top 10 percent of their classes, in which case the FSEE is not required for appointment. The United States Civil Service Commission in Washington has asked us to test employees after entrance on duty--this is for research purposes, I understand."

TABLE A-1

REGIONAL SURVEY: EXTENT AND KIND OF USE OF THE FSEE

| | Agency Regional Offices | | | | CSC | CSC | |
|---|---|---|---|---|---|---|---|
| | OEO | HEW | DOL | HUD | Regional Offices | Area Offices: (Job Information Centers) | Total |
| SENT | 10 | 10 | 10 | 10 | 10 | 53 | 103 |
| RECEIVED | 9 | 2 | 4 | 3 | 7 | 7 | 32 |
| Data | 8 | 2 | 0 | 2 | 0 | 0 | 12 |
| No Data | 1 | 0 | 4 | 1 | 7 | 7 | 20 |
| % RETURN | 90% | 20% | 40% | 30% | 70% | 13% | 31% |

APPENDIX B

THE URBAN INSTITUTE 2100 M STREET, N.W. WASHINGTON, D.C. 20037 (202) 223-1950

Dear Mr.

The Urban Institute is presently undertaking a study of applicant examinations and evaluations which are used by the Civil Service Commission in determining suitability for Federal employment. We are particularly interested in the role of the Federal Service Entrance Examination in these evaluations and the validation procedures which have been developed and are being used for this examination.

We understand that the use of this examination in evaluating federal employment candidates will be a subject of litigation in the near future. It is our intention to make the results of this research available to both the litigants and to the court in the hope that it will aid in the court's decisions.

In conducting this study we are trying to develop as broad a base of information as possible. Consequently, we would greatly appreciate your attention to the accompanying questionnaire. We realize that this will take some valuable time. However, since we hope assembling these data will be of great importance in helping the court to reach a wise decision in this far-reaching case, we would be most grateful for your help.

Sincerely,

Worth Bateman
Vice President for Operations

Enclosures

APPENDIX C

THE URBAN INSTITUTE 2100 M STREET, N.W. WASHINGTON, D.C. 20037 (202) 223-1950

PUBLIC INTEREST STUDY
PROJECT # 712 - 2
FEDERAL SERVICE ENTRANCE EXAMINATION

## REGIONAL SURVEY

Agency: _____ Respondent: Name: _____

_____ Title: _____

Region: _____ Address: _____

_____

1. To what extent are personnel in Grade levels GS-5 through GS-9 initially drawn from the FSEE register?
   a. exclusively from register (100%) _____
   b. mostly from register (70-99%) _____
   c. about half from register, half from agency recruitment (30-69%) _____
   d. mostly from agency recruitment (1-29%) _____
   e. exclusively from agency recruitment (0%) _____

2. Please list other factors beside the FSEE which are used in evaluation of these candidates.

   _____ _____

   _____ _____

   _____ _____

   _____ _____

3. Does the weight given the FSEE in the total evaluation
   a. vary with the applicant? Yes ____ No ____
   b. vary with the position? Yes ____ No ____
   c. vary with the grade level? Yes ____ No ____

   If yes for any of the above, please explain in what ways they vary.

   _____

   _____

4. On the average, what weight or proportion of the total evaluation does the FSEE hold? (Please give approximate percent.) ____%

5. a. Does your examining or regional office have both biographical data and FSEE scores for each candidate broken down by FSEE testing site?

Yes, both _____
Just FSEE scores _____
Neither _____

b. If yes, do these data contain minority group and/or ethnic information?

Yes _____ No _____

6. What measures do you employ as indicators of job performance?

_____

_____

_____

7. Of those made eligible by the FSEE and presently working:

a. are data available on their FSEE scores? Yes _____ No _____
(1) by minority and/or ethnic group? Yes _____ No _____
b. are data available on job performance ratings? Yes _____ No _____
(1) by minority and/or ethnic group? Yes _____ No _____
c. are data available on promotion action taken? Yes _____ No _____
(1) by minority and/or ethnic group? Yes _____ No _____
d. others_____ Yes _____ No _____

_____

8. Has your office requested or participated in any study of the validity of the FSEE? Yes _____ No _____

If yes, please answer the following questions:

a. Did you feel the validation study indicated that the FSEE was predictive of:
(1) successful on-the-job performance? Yes _____ No _____
(2) trainability for higher positions? Yes _____ No _____
(3) other (list)_____

_____

b. Did the study indicate the FSEE had predictive value for different population groups? Yes _____ No _____
If yes, please indicate in what manner._____

_____

_____

c. Have the study results been published? Yes _____ No _____
If yes, please send the Urban Institute a copy of the report(s) or indicate below where the report(s) can be obtained.

d. If the study results have not been published, please describe briefly the purpose, scope, procedure and principal findings of the study.

_____

_____

_____

THANK YOU. YOUR ATTENTION AND COOPERATION ARE APPRECIATED.